

Abstract

Natural-sounding synthesized speech is obtained from pieced elemental speech units that have their super-class identities known (e.g. phoneme type), and their line spectral frequencies (LSF) set in accordance with a correlation between the desired fundamental frequency and the LSF vectors that are known for different classes in the super-class. The correlation between a fundamental frequency in a class and the corresponding LSF is obtained by, for example, analyzing the database of recorded speech of a person and, more particularly, by analyzing frames of the speech signal.